US 20200249751A1

(19) **United States**

(12) **Patent Application Publication** (10) **Pub. No.: US 2020/0249751 A1**

Norden (43) **Pub. Date: Aug. 6, 2020**

(54) **USER INPUT PROCESSING WITH EYE TRACKING**

(71) Applicant: **SONY INTERACTIVE ENTERTAINMENT AMERICA LLC**, San Mateo, CA (US)

(72) Inventor: **Chris Norden**, Austin, TX (US)

(21) Appl. No.: **16/701,832**

(22) Filed: **Dec. 3, 2019**

Related U.S. Application Data

(63) Continuation of application No. 15/296,552, filed on Oct. 18, 2016, now Pat. No. 10,496,159, which is a continuation of application No. 13/464,703, filed on May 4, 2012, now Pat. No. 9,471,763.

Publication Classification

(51) **Int. Cl.**

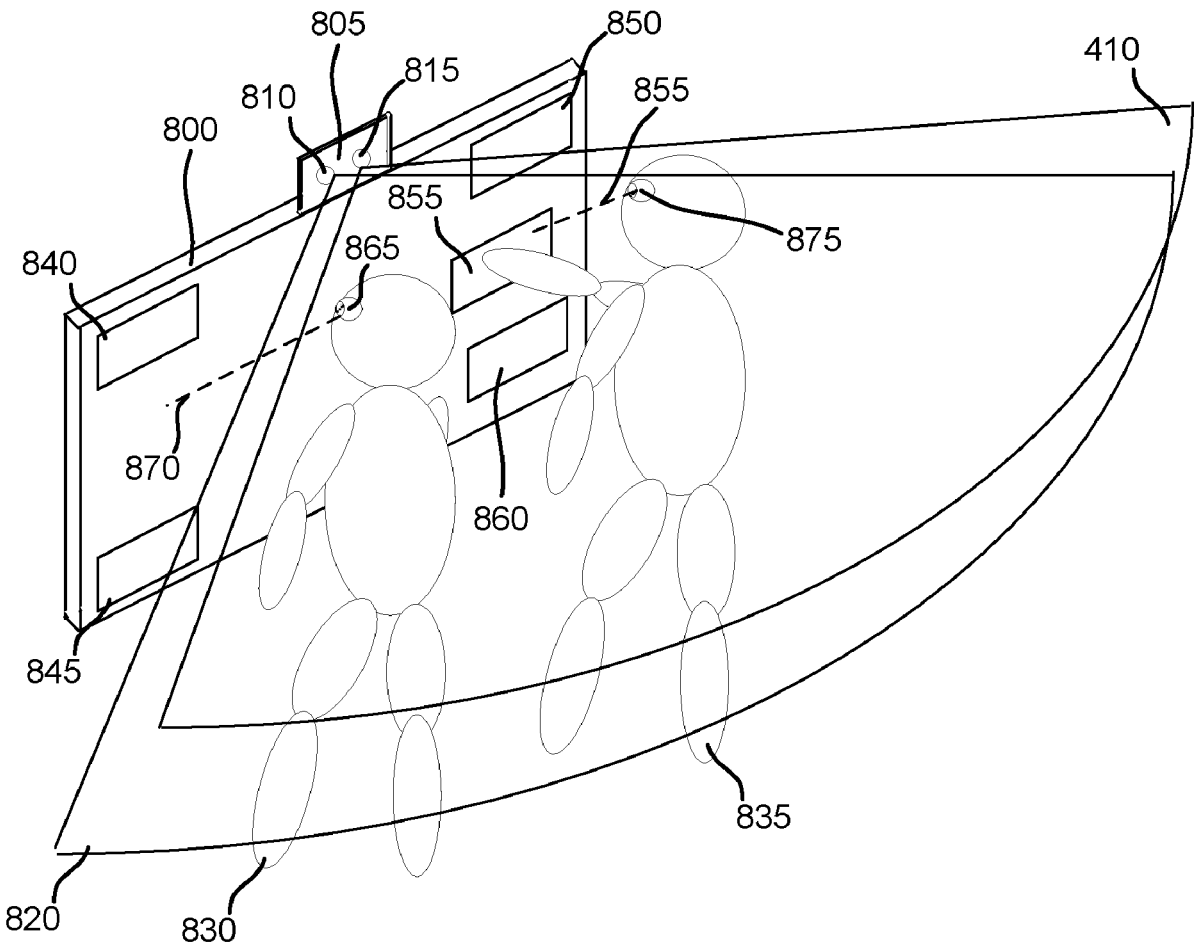
G06F 3/01	(2006.01)
G06F 3/041	(2006.01)
G06F 21/32	(2006.01)
G06K 9/00	(2006.01)
G06F 3/0488	(2006.01)

(52) **U.S. Cl.**

CPC ***G06F 3/013*** (2013.01); ***G06F 3/041*** (2013.01); *G06F 2221/2133* (2013.01); ***G06K 9/00597*** (2013.01); ***G06F 3/0488*** (2013.01); ***G06F 21/32*** (2013.01)

(57) **ABSTRACT**

A system determines which user of multiple users provided input through a single input device. A mechanism captures images of the one or more users. When input is detected, the images may be processed to determine which user provided an input using the input device. The images may be processed to identify each users head and eyes, and determine the focus point for each user's eyes. The user which has eyes focused at the input location is identified as providing the input. When the input mechanism is a touch screen, the user having eyes focused on the touch screen portion which was touched is identified as the source of the input.



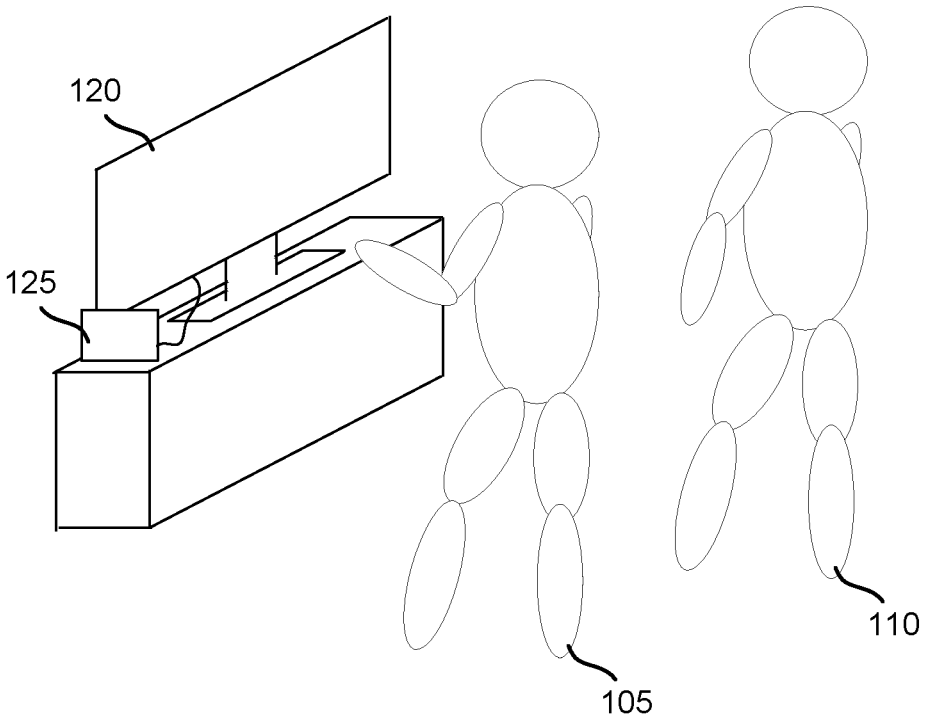


FIGURE 1

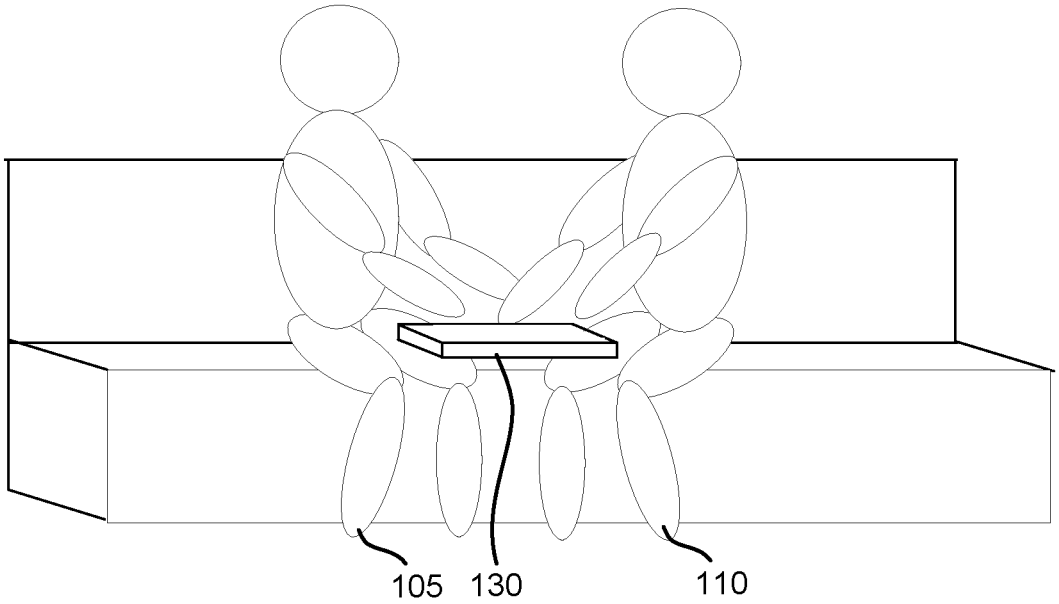


FIGURE 2

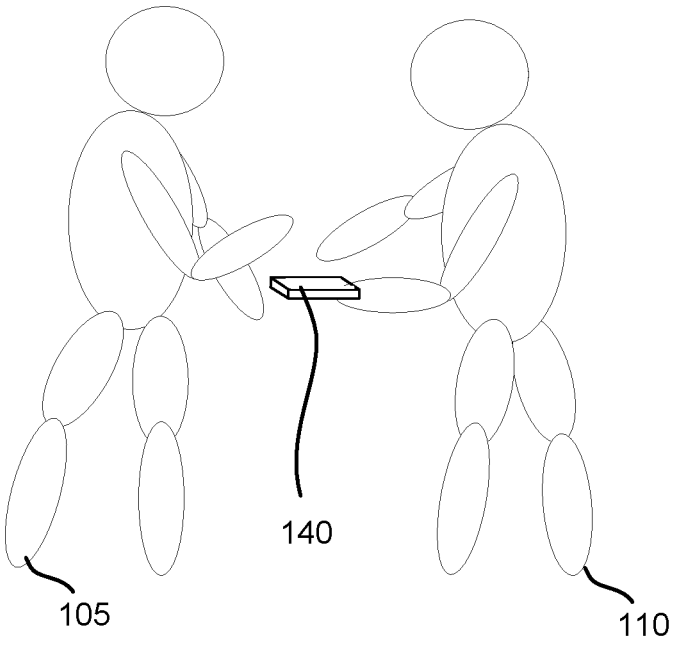


FIGURE 3

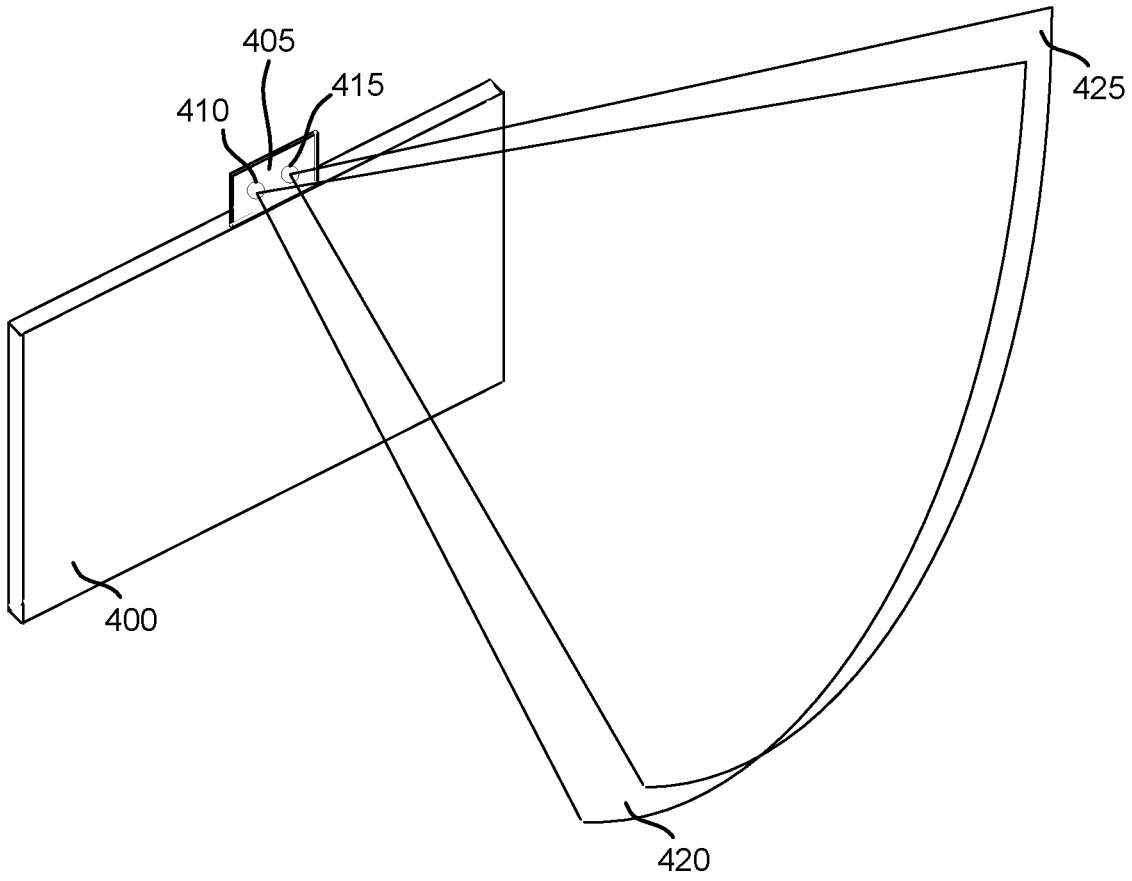


FIGURE 4

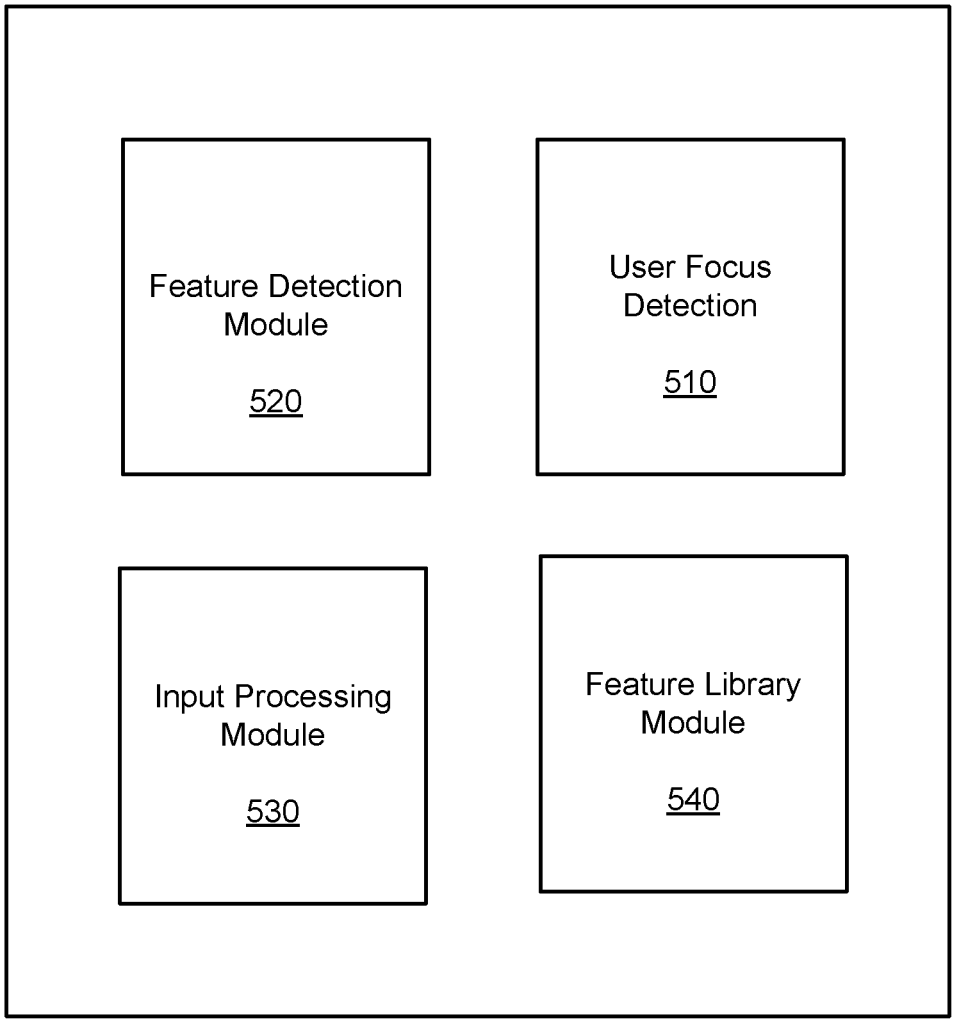

Feature Detection Module
520
User Focus Detection
510
Input Processing Module
530
Feature Library Module
540

FIGURE 5

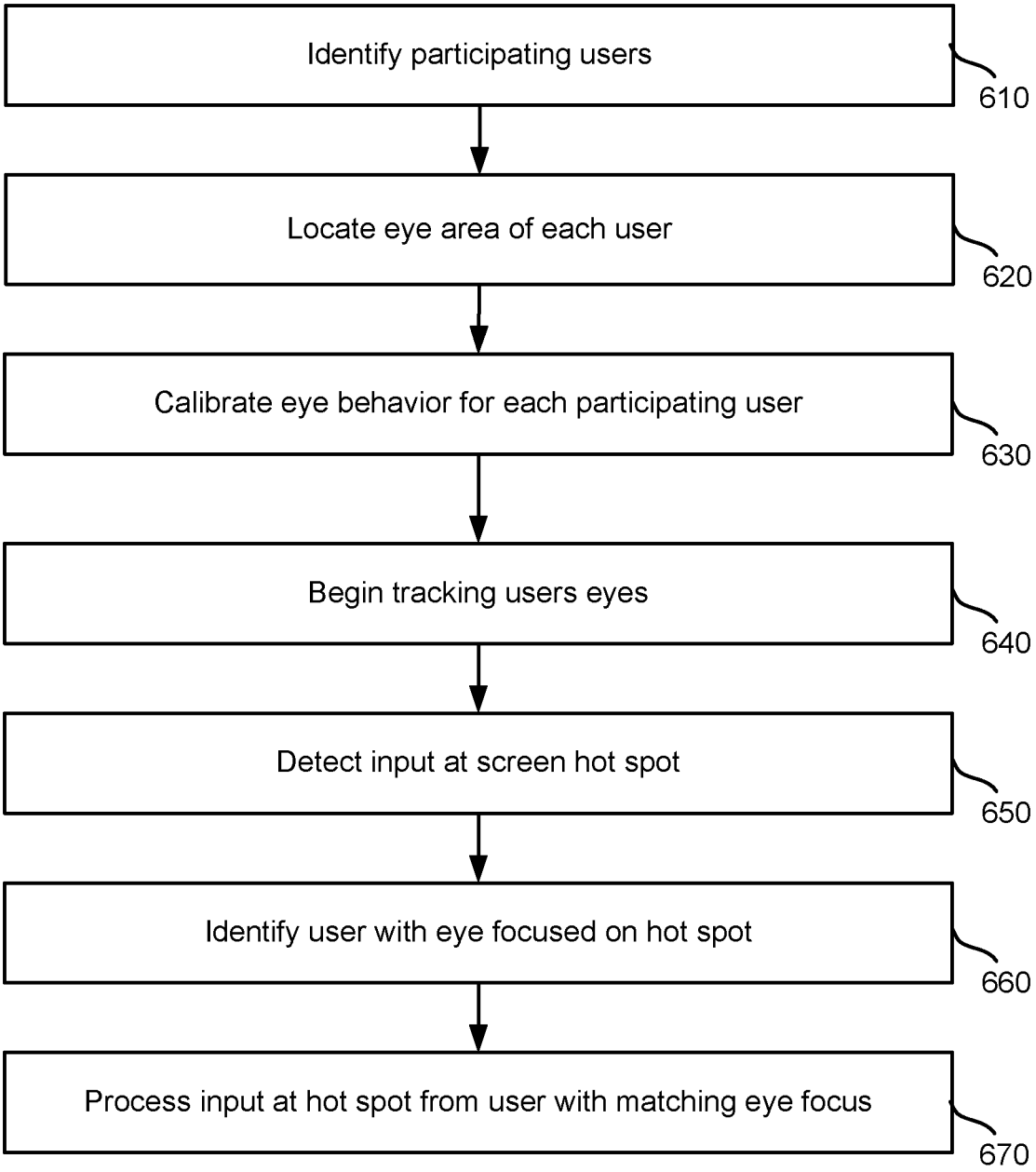

Identify participating users
610
Locate eye area of each user
620
Calibrate eye behavior for each participating user
630
Begin tracking users eyes
640
Detect input at screen hot spot
650
Identify user with eye focused on hot spot
660
Process input at hot spot from user with matching eye focus
670

FIGURE 6

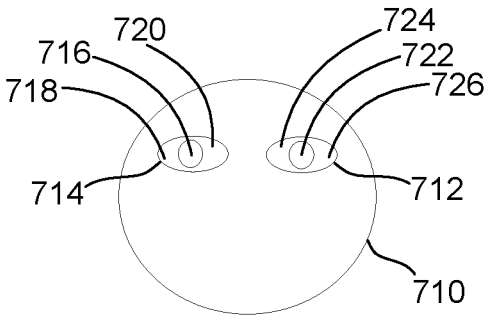


FIGURE 7A

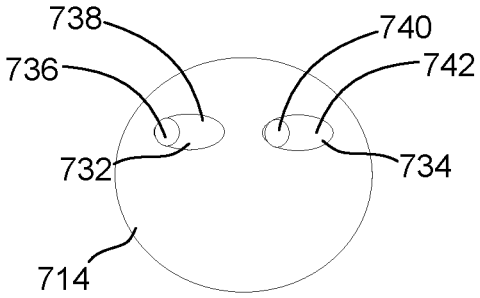


FIGURE 7B

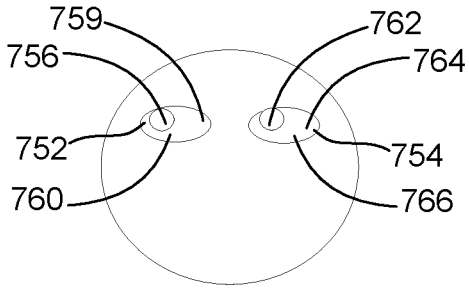


FIGURE 7C

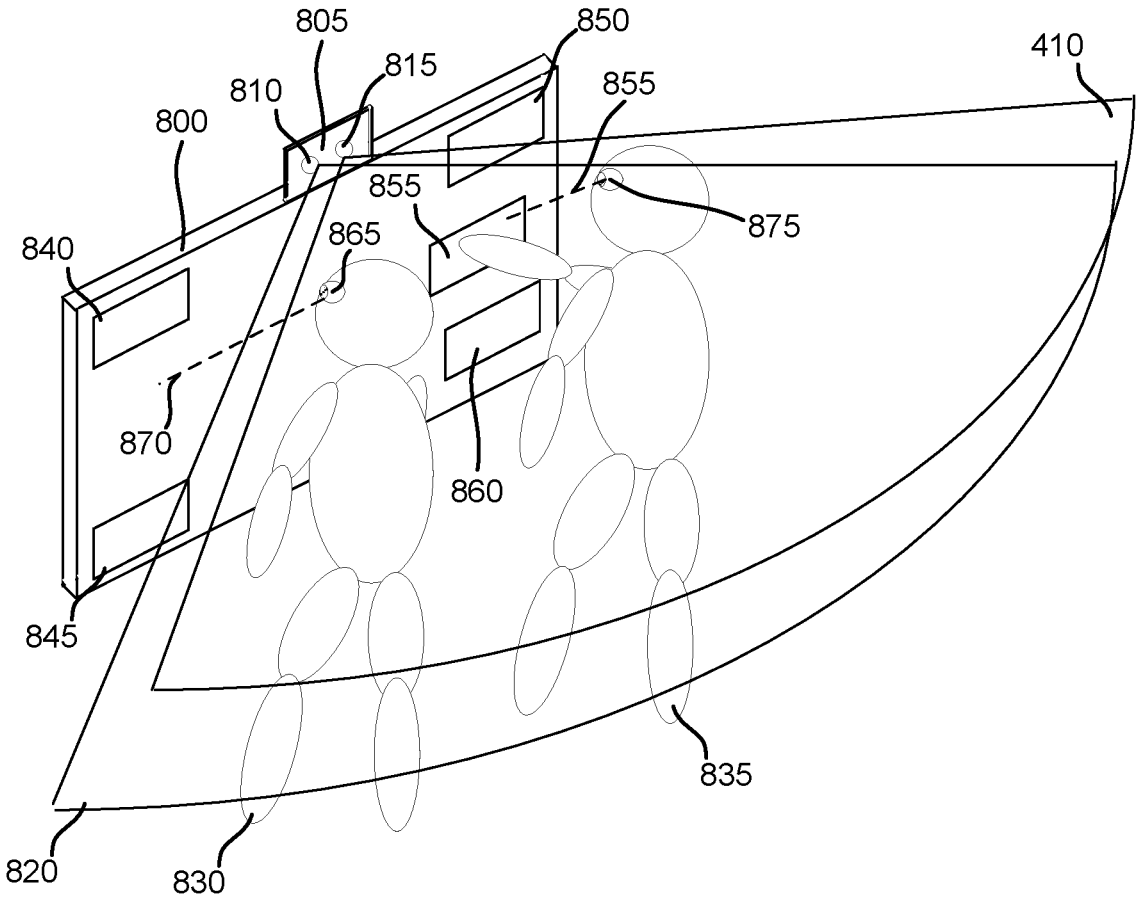

805
850
410
815
810
855
800
855
875
840
865
870
860
845
835
820
830

FIGURE 8

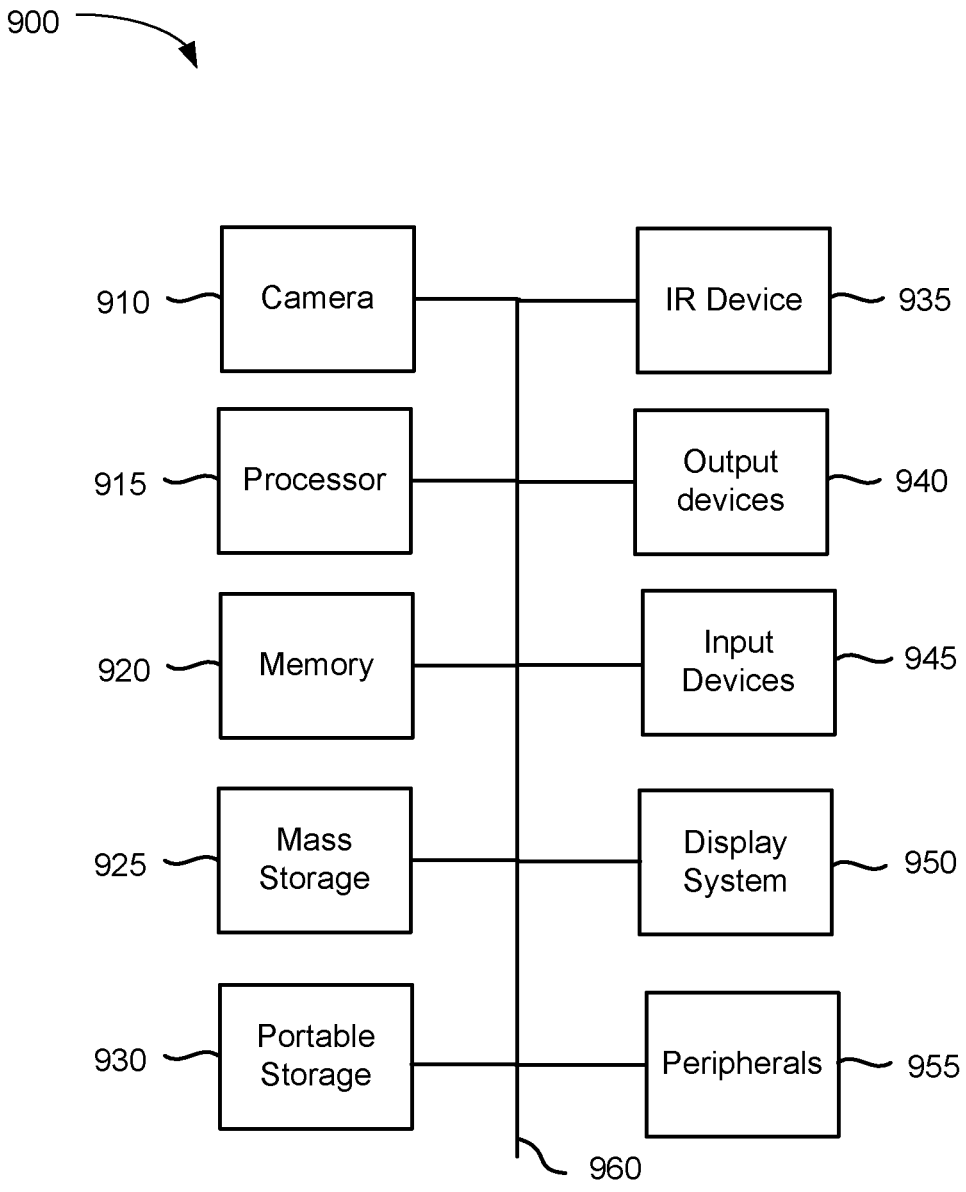

900
910
Camera
IR Device
935
915
Processor
Output devices
940
920
Memory
Input Devices
945
925
Mass Storage
Display System
950
930
Portable Storage
Peripherals
955
960

FIGURE 9

USER INPUT PROCESSING WITH EYE TRACKING

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application is a continuation and claims the priority benefit of U.S. patent application Ser. No. 15/296,552 filed Oct. 18, 2016, now U.S. Pat. No. 10,496,159, which is a continuation and claims the priority benefit of U.S. patent application Ser. No. 13/464,703 filed May 4, 2012, now U.S. Pat. No. 9,471,763, the disclosures of which are incorporated herein by reference.

BACKGROUND OF INVENTION

1. Field of the Invention

[0002] The present invention generally relates to input processing systems and more particularly to processing input from a plurality of users via image processing.

2. Description of the Prior Art

[0003] Computing system applications interact with users by receiving user input, processing the input, and providing a result. As systems become more advanced and more people embrace technology, applications have evolved to engage multiple users simultaneously. Multiple users may interact with computing systems at the same time, and the users may provide output that relates to a particular user. For example, two users may use game controllers to play a computer game that allows the users to compete against each other during a game session. Each user provides input with a respective gaming controller.

[0004] Technology has evolved to allow different types of interaction with computing systems. Rather than receiving input from a remote gaming controller for each user for example, some computing systems utilize a single input mechanism such as a touch screen. When only one user is engaging with the computing device, the single input mechanism receives input from the single user. When multiple users engage the computing system with a single input mechanism, it is very difficult to determine which user is providing the input.

[0005] There is a need in the art for a system that allows multiple users to easily and efficiently interact with a computing device using a single input mechanism.

SUMMARY OF THE CLAIMED INVENTION

[0006] In an embodiment, performance, functionality, content, or business relevancy. Based on learning techniques, efficient monitoring, and resource management, the present system may capture data for and provide analysis information for outliers of a web application with very low overhead.

[0007] In an embodiment, input may be received by first identifying a plurality of users physically in the presence of the device. An input may be received by the device from a first user of the plurality of physically present users. A physical state may be detected from one of the plurality of users associated with the input.

[0008] In an embodiment, a system for detecting input may include a display device, a camera, a processor and modules stored in memory and executable by the processor. The camera may capture color image data and provide the image data to the processor. A feature detection module is executable to detect a physical feature of a user. A user focus detection module detects the point of focus of a user's eyes. An input processing module receives and processes an input from a user.

BRIEF DESCRIPTION OF THE DRAWINGS

[0009] FIG. **1** is a block diagram of an exemplary system for receiving input through a touch screen of a computing console.

[0010] FIG. **2** is a block diagram of an exemplary system for receiving input through a touch screen of a tablet computer.

[0011] FIG. **3** is a block diagram of an exemplary system for receiving input through a touch screen of a mobile device.

[0012] FIG. **4** is an exemplary computing device with user detection mechanism.

[0013] FIG. **5** is an exemplary set of executable software modules.

[0014] FIG. **6** is a flow chart of an exemplary method for identifying input from one of a plurality of users.

[0015] FIG. **7**A-FIG. **7**C are exemplary images of a user's eyes.

[0016] FIG. **8** is a block diagram of an exemplary computing device receiving input from one of multiple users.

[0017] FIG. **9** is a block diagram of an exemplary system for implementing a computing device.

DETAILED DESCRIPTION

[0018] Embodiments of the invention determine which user of multiple users provided input through a single input device. The computing system may include a mechanism for capturing images of the one or more users. The images may be processed to determine which user provided an input using the input device. For example, the images may be processed to identify each users head and eyes, and determine the focus point for each user's eyes. The user which has eyes focused at the input device is identified as providing the input. In embodiments where the input mechanism may be a touch screen, the user having eyes focused on the touch screen portion which was touched is identified to be providing the input.

[0019] Embodiments of the invention may be used with several types of computing devices. FIG. **1** is a block diagram of an exemplary system for receiving input through a touch screen of a computing console. A computing console **125** provides image data for display to touch screen **120**. Touch screen **120** may receive touch input and provide the input computing console **125**. The computing console may receive the input, process the input and create new image data, and provide the image data to touch screen **120**. Any of users **105** and **110** may provide input to computing console **125**, for example by touching touch screen **120**. The present invention is able to determine which user touched the touch screen and process the input accordingly. FIG. **2** is a block diagram of an exemplary system for receiving input through a touch screen of a tablet computer **130**. Users **105** and **110** may each provide input to tablet computer **130** using a touch screen of the tablet. Embodiments of the invention may determine which of users **105** and **110** provided input to the tablet computer **130** and process the input accordingly. FIG. **3** is a block diagram of an exemplary system for

receiving input through a touch screen of a mobile device. Users **105** and **110** may each provide input to the mobile device **140** through touch screen or other input. The present invention may determine which user provided input to mobile device **140** and process the input for that user.

[0020] FIG. **4** is an exemplary computing device with user detection mechanism **45**. User detection mechanism **405** may be used with a computing console, tablet compute, mobile device, or any other computing device. User detection mechanism includes color camera **410** and infra-red (IR) camera **415**. Color camera **410** may capture images of an area **420** in the vicinity of the computing device. The images may be processed to identify users, physical features of users, and a state of the physical features. For example, images of two players captured by color camera **410** may be processed to identify two users within area **420**, each user's physical features including user head and eyes, and the state of each user's eyes such as where each user is focusing their eyes. An IR imaging system **415** may also be used to capture and process images. The IR imaging system **415** may be utilized in low light conditions to capture IR images of area **425**. The IR images may be processed to identify a number of users, user physical features, and feature states similar to the processing of color camera image processing.

[0021] FIG. **5** is an exemplary set of executable software modules. The modules may be executed by a processor to implement various aspects of the invention described herein. A feature detection module **520** detects features in images of a user. For example, the feature detection module may receive an image, detect contours of a human face, and identify eyes within a face. Module **520** may have template of one or more faces to compare to portions of received images to identify a human face. The template may be stored in feature library module **540**. In some embodiments, feature detection module **540** may detect motion between two or more consecutive images and use the change in pixel coloration to detect users. Once a user head is detected, the feature detection module analyzes features within the head to detect user's eyes. The eyes may be identified by detecting features such as a nose which are close to the eyes, determining that the user eyes are a certain distance below the top of the user's head, or perform other algorithms that are generally known in the art. In addition to physical features, other features may be determined as well, such as for example the distance a user is away from the computing device. In some embodiments, if a detected person is greater that a certain threshold distance from the computing device, the user will not be considered a user capable of providing input.

[0022] User focus module **510** may analyze images of a user's eye to determine where the user is focusing. The front of a human eye includes a black pupil, a colored iris around the pupil, and a white sclera around the iris. A computing device may analyze the area and location of the sclera to determine if a user is focused up, down, left and right. For example, when a user's eyes are focused on an object to his right, an image captured of the user's eyes will show more of the user's sclera on the right side of the eye in the image (the user's left side) than the left side, because the eyes will be moved towards the left side.

[0023] Input focusing module **530** receives input and processes the input. The input may be selection of a designated hot spot on a touch screen, a button, a wireless signal, or some other input. Input focusing module **530** may receive information from other modules which identify a user that has provided a recent input. The input processing module then processes the input as the identified user's action.

[0024] Feature library module **540** may include facial and eye masks, templates, models and other data used to process an image and identify a user physical feature and the feature state, such as which direction a user's eyes are focused at.

[0025] FIG. **6** is a flow chart of an exemplary method for identifying input from one of a plurality of users. The method of FIG. **6** may be performed by any of computing devices **120** and **125**, **130** and **140**. Participating users are identified by the computing device at step **610**. The participating users are those that may provide input to the computing device. The participating users may be identified by registration with the system or image processing. Registration may include each user providing identification information of some sort, or an indication that they are present, to the computing device. Image processing may include capturing one or more images of an area from which users may provide input from, detecting the number of human heads, and assigning a value to each detected head. In some embodiments, both registration and image processing may be used to identify participating users.

[0026] A method for detecting a human head via image processing may begin with analyzing an image for shapes resembling a human head. Shapes may be identified using contrast detection, motion detection, and other techniques. Once a potential head shape is detected, the head candidate is analyzed for features common to most human heads. The features may include contrast, shading or other features present where a nose, mouth, or eyes may be. If the candidate head satisfies a threshold level of features, the head candidate may be identified as a participating user. Other methods for detecting faces in images are known in the art.

[0027] An eye area of each user is located at step **615**. Detecting an eye area of a participating user's head may involve searching for a contrast, brightness, or other image property level at about the area within the head that the eye is located. Once the user eyes are located, eye behavior may be calibrated for each participating user at step **620**. Calibration may include instructions on the screen to a participating user indicating a distance range from the computing device the user's face should, instructions to look at a particular point on the screen, and other directions. The calibration may have a user look at different points or hot spots on the display of the computing device, and analyze the images of the user's eyes when the focus of the user's eyes is known. For example, FIG. **7**A-FIG. **7**C illustrate a user head and eyes as the user eyes are focused on different areas. FIG. **7**A illustrates a user head **710** with eyes **714** and **712**. Eye **714** includes iris and pupil area **716** and sclera areas **718** and **720**. Eye **712** includes iris and pupil area **722** and sclera areas **724** and **726**. When the sclera areas to the left and right of the iris and pupil areas are about the same, the user may be determined to be focused about straight ahead. When the area of the sclera is much larger to the right of the iris and pupil area than the sclera area to the left of the pupil and iris, the user focus may be towards the right of the user (FIG. **7**B). Similarly, when the area of the sclera is much larger to the right and underneath the iris and pupil area than the sclera area to the left and above the pupil and iris, the user focus may be towards the upper right of the user (FIG. **7**C). The degrees of focus and corresponding focus

point may be derived from taking a set of measurements of the user's eyes during calibration. The area and location of the user's pupil, sclera and other objects may be recorded for subsequently tracking user eyes.

[0028] Tracking of user eyes begins at step **625**. The tracking involves capturing consecutive images of the user. The images may be processed to track and maintain knowledge of the user eye location and focus. In some embodiments, the images are captured repeatedly and stored, but are then discarded if not input is received from any user.

[0029] Input is detected at a screen hot spot at step **630**. The input may include a user touching a hot spot on a touch screen for one of computing console **120**, tablet computer **220**, and mobile device **320**. The hot spot may be a particular image object displayed on the screen, such as an image of a character, ball, virtual item, text, or other object. A user having eyes focused on the hot spot location of the input is identified at step **635**. The eye location may be determined as discussed above with respect to the calibration process. A user corresponding to the particular input may be identified in a variety of ways. For example, the eye focus for each user may be determined, and the user with the eye focus closest to the hot spot may be selected. Alternatively, the eye focus for each user may be determined until an eye focus is detected within a threshold distance of the hot spot at which input was received. In some embodiments, a likelihood of input may be determined for each user based on their input history, their eye focus, whether an input is expected from the user, and so on. Once an input is associated with an identified user, the input at the hot spot is processed for the particular user at step **640**.

[0030] FIG. **8** is a block diagram of an exemplary computing device receiving input from one of multiple users. Computing device **800** includes a camera system **805** (user detection system) with color camera **810** and IR camera **815**. Users **830** and **835** are each in the range **820** of the color camera **810** and range **825** of the IR camera **815**. The computing device **800** includes a touch screen having hot spots **840**, **845**, **850**, **855**, and **860**. During calibration, the system may display text that asks the user to focus on a particular hot spot while the system captures an image of the user eyes.

[0031] When input is received, the computing device **800** determines the user focus and identifies the input as coming from the user focusing on the hot spot that received the input. For example, user **835** has provided input at hot spot **855** by pressing the screen at hot stop **855**. Upon receiving the input, the processing device will analyze images captured from the color camera **810**, IR camera **815**, or both. From the images, the users focus will be determined. If after processing the images, user **830** is determined to have a focus of **870** and user **835** is determined to have focus **880** which corresponds to hot step **855**, the input received at hot spot **855** will be associated with user **835**. By determining where user eyes are focused, an input received through a device used by a plurality of players may be associated with one of a plurality of players.

[0032] FIG. **9** illustrates an exemplary computing system **900** that may be used to implement a computing device for use with the present technology. System **900** of FIG. **9** may be implemented in the contexts of the likes computing console **120**, tablet computer **220**, and mobile device **320**. The computing system **900** of FIG. **9** includes one or more cameras **910**, processors **915** and memory **920**. Main memory **920** stores, in part, instructions and data for execution by processor **915**. Main memory **920** can store the executable code when in operation. The system **900** of FIG. **9** further includes a mass storage device **925**, portable storage medium drive(s) **930**, IR device **935**, output devices **940**, user input devices **945**, a display **950**, and peripheral devices **955**.

[0033] Cameras **910** may include one more cameras able to capture a series of photos suitable for image processing analysis. The photos may be embedded within the computing system of mounted externally to the system. The images captured by camera **910** may be provided to processor **915** via bus **960**, which may execute modules stored in memory **920** to analyze the images for features detection.

[0034] IR device **935** may include an IR camera that is able to capture images in very low light conditions. The IR images may be processed similarly as color camera images for user feature detection. The images captured from IR device **935** may be sent to processor **915** for processing via bus **960**.

[0035] The components shown in FIG. **9** are depicted as being connected via a single bus **990**. However, the components may be connected through one or more data transport means. For example, processor unit **915** and main memory **920** may be connected via a local microprocessor bus, and the mass storage device **925**, peripheral device(s) **955**, portable storage device **930**, and display system **950** may be connected via one or more input/output (I/O) buses.

[0036] Mass storage device **925**, which may be implemented with a magnetic disk drive or an optical disk drive, is a non-volatile storage device for storing data and instructions for use by processor unit **915**. Mass storage device **925** can store the system software for implementing embodiments of the present invention for purposes of loading that software into main memory **920**.

[0037] Portable storage device **930** operates in conjunction with a portable non-volatile storage medium, such as a floppy disk, compact disk or Digital video disc, to input and output data and code to and from the computer system **900** of FIG. **9**. The system software for implementing embodiments of the present invention may be stored on such a portable medium and input to the computer system **900** via the portable storage device **930**.

[0038] Input devices **945** provide a portion of a user interface. Input devices **945** may include an alpha-numeric keypad, such as a keyboard, for inputting alpha-numeric and other information, or a pointing device, such as a mouse, a trackball, stylus, or cursor direction keys. Additionally, the system **900** as shown in FIG. **9** includes output devices **940**. Examples of suitable output devices include speakers, printers, network interfaces, and monitors.

[0039] Display system **950** may include a liquid crystal display (LCD) or other suitable display device. Display system **950** receives textual and graphical information, and processes the information for output to the display device. Display system **950** may include a touch screen device which receives input by detecting a touch on the surface of the display. The pixels receiving the touch are communicated to processor **915** via bus **960**.

[0040] Peripherals **955** may include any type of computer support device to add additional functionality to the computer system. For example, peripheral device(s) **955** may include a modem or a router.

[0041] The components contained in the computer system **900** of FIG. **9** are those typically found in computer systems that may be suitable for use with embodiments of the present invention and are intended to represent a broad category of such computer components that are well known in the art. Thus, the computer system **900** of FIG. **9** can be a personal computer, hand held computing device, telephone, mobile computing device, workstation, server, minicomputer, mainframe computer, or any other computing device. The computer can also include different bus configurations, networked platforms, multi-processor platforms, etc. Various operating systems can be used including Unix, Linux, Windows, Macintosh OS, Palm OS, and other suitable operating systems.

[0042] The foregoing detailed description of the technology herein has been presented forpurposes of illustration and description. It is not intended to be exhaustive or to limit the technology to the precise form disclosed. Many modifications and variations are possible in light of the above teaching. The described embodiments were chosen in order to best explain the principles of the technology and its practical application to thereby enable others skilled in the art to best utilize the technology in various embodiments and with various modifications as are suited to the particular use contemplated. It is intended that the scope of the technology be defined by the claims appended hereto.

1. (canceled)

2. A method for identifying input based on eye focus, the method comprising:

capturing image data for a physical area proximate to a computing device, the image data captured by a camera;

evaluating the captured image data to identify one or more eyes of a user present in the physical area;

tracking movement of a focal point of the eyes, wherein tracking the movement is based on image data captured by the camera;

generating a display of a virtual item associated with a predefined input point of the display;

identifying when the tracked focal point moves to a location of the virtual item within the generated display; and

activating the predefined input point associated with the virtual item, wherein activating the predefined input point includes displaying new data at the location of the virtual item within the display.

3. The method of claim **2**, further comprising identifying that the user has touched the virtual item to activate the predefined input point.

4. The method of claim **2**, further comprising:

displaying a first virtual item at a first location on the display when calibrating the movement tracking;

identifying that a first focus point of the user eyes corresponds to the first location of the first virtual item;

displaying the first virtual item at a second location on the display; and

completing the calibration of the movement tracking after identifying that a second focus point of the user eyes corresponds to the second location on the display.

5. The method of claim **2**, further comprising:

capturing an initial image of the user eyes; and

analyzing the initial captured image to identify a feature of the user eyes.

6. The method of claim **5**, further comprising identifying a location of the feature based on the analysis.

7. The method of claim **6**, wherein the feature is an iris and the movement tracking is based at least in part on the identified location of the iris.

8. The method of claim **2**, further comprising:

identifying an object that is a distance away from the computing device; and

identifying the distance that the object is away from the computing device based on an analysis of the captured image data.

9. The method of claim **2**, further comprising:

identifying a set of eyes of a second user; and

calibrating eye motion of the set of eyes of the second user by:

displaying a first virtual item at a first location on the display,

identifying that a first focus point of the set of eyes corresponds to the first location of the first virtual item,

displaying the first virtual item at a second location on the display, and

completing the calibration the eye motion of the second user after identifying that a second focus point of the one or more eyes corresponds to the second location on the display.

10. The method of claim **2**, further comprising processing color data collected by a color camera at the apparatus.

11. The method of claim **2**, wherein the camera is an infrared (IR) camera.

12. A non-transitory computer-readable storage medium having embodied thereon a program executable by a processor to implement a method for identifying input based on eye focus, the method comprising:

capturing image data for a physical area proximate to a computing device, the image data captured by a camera;

evaluating the captured image data to identify one or more eyes of a user present in the physical area;

tracking movement of a focal point of the eyes, wherein tracking the movement is based on image data captured by the camera;

generating a display of a virtual item associated with a predefined input point of the display;

identifying when the tracked focal point moves to a location of the virtual item within the generated display; and

activating the predefined input point associated with the virtual item, wherein activating the predefined input point includes displaying new data at the location of the virtual item within the display.

13. The non-transitory computer-readable storage medium of claim **12**, the program further executable to identify that the user has touched the virtual item to activate the predefined input point.

14. The non-transitory computer-readable storage medium of claim **12**, the program further executable to:

display a first virtual item at a first location on the display when calibrating the movement tracking;

identify that a first focus point of the user eyes corresponds to the first location of the first virtual item;

display the first virtual item at a second location on the display; and

complete the calibration of the movement tracking after identifying that a second focus point of the user eyes corresponds to the second location on the display.

15. The non-transitory computer-readable storage medium of claim **12**, the program further executable to:

capture an initial image of the user eyes; and

analyze the initial captured image to identify a feature of the user eyes.

16. The non-transitory computer-readable storage medium of claim **15**, the program further executable to identify a location of the feature based on the analysis.

17. The non-transitory computer-readable storage medium of claim **16**, wherein the feature is an iris and the movement tracking is based at least in part on the identified location of the iris.

18. The non-transitory computer-readable storage medium of claim **12**, the program further executable to:

identify an object that is a distance away from the computing device; and

identify the distance that the object is away from the computing device based on an analysis of the captured image data.

19. The non-transitory computer-readable storage medium of claim **12**, the program further executable to:

identify a set of eyes of a second user; and

calibrate eye motion of the set of eyes of the second user by:

displaying a first virtual item at a first location on the display,

identifying that a first focus point of the set of eyes corresponds to the first location of the first virtual item,

displaying the first virtual item at a second location on the display, and

completing the calibration the eye motion of the second user after identifying that a second focus point of the one or more eyes corresponds to the second location on the display.

20. The non-transitory computer-readable storage medium of claim **2** the program further executable to process color data collected by a color camera at the apparatus.

21. A system for identifying input based on eye focus, the system comprising:

a camera that captures image data for a physical area proximate to a computing device;

a memory;

a processor that executes instructions out of the memory to:

evaluate the captured image data to identify one or more eyes of a user present in the physical area, and

track movement of a focal point of the eyes, wherein tracking the movement is based on image data captured by the camera; and

a display that displays a virtual item associated with a predefined input point of the display, wherein the processor also executes instructions out of the memory to:

identify when the tracked focal point moves to a location of the virtual item within the display, and

activate the predefined input point associated with the virtual item, the activating of the predefined input point including displaying new data at the location of the virtual item within the display.

* * * * *